GEOPHYSICAL RESEARCH LETTERS, VOL. 40, 3048–3054, doi:10.1002/grl.50615, 2013

# Discovery and analysis of topographic features using learning algorithms: A seamount case study

Andrew P. Valentine,[1] Lara M. Kalnins,[2] and Jeannot Trampert[1]

Received 6 May 2013; revised 31 May 2013; accepted 31 May 2013; published 20 June 2013.

[1] Identifying and cataloging occurrences of particular topographic features are important but time-consuming tasks. Typically, automation is challenging, as simple models do not fully describe the complexities of natural features. We propose a new approach, where a particular class of neural network (the "autoencoder") is used to assimilate the characteristics of the feature to be cataloged, and then applied to a systematic search for new examples. To demonstrate the feasibility of this method, we construct a network that may be used to find seamounts in global bathymetric data. We show results for two test regions, which compare favorably with results from traditional algorithms. **Citation:** Valentine, A. P., L. M. Kalnins, and J. Trampert (2013), Discovery and analysis of topographic features using learning algorithms: A seamount case study, *Geophys. Res. Lett.*, *40*, 3048–3054, doi:10.1002/grl.50615.

## 1. Introduction

[2] Often in the geosciences, it is necessary to identify particular features in topographic data—perhaps a certain type of landform. Typically, an experienced researcher can readily recognize the characteristics of the feature, and for a small region, hand-selection may be straightforward. However, to compile a regional or global catalog, automation becomes necessary and may bring benefits such as repeatability and objectivity. Unfortunately, most features of interest are not adequately described by simple mathematical models, so designing appropriate classification systems is challenging.

[3] As a concrete example, we consider the problem of cataloging seamounts—isolated topographic highs of volcanic origin—in the oceans. Such a catalog is important for understanding tectonic and oceanic processes and also has relevance in fields such as marine ecology and submarine navigation [see also *Kim and Wessel*, 2011]. The earliest studies [e.g., *Menard*, 1959; *Smith and Jordan*, 1988] used ship track bathymetry data with extremely limited coverage; with the advent of satellite-derived altimetry and gravity measurements, global studies became possible [e.g., *Craig and Sandwell*, 1988]. Typically, work has focused on detecting the gravitational signature of seamounts in the vertical gravity gradient (VGG) [e.g., *Wessel and Lyons*, 1997; *Kim and Wessel*, 2011]. Alternative strategies include those of *Hillier* [2008], based on wavelets, and *Kitchingman and Lai* [2004] and *Morato et al.* [2008], where topography is inverted and tools developed for flow modeling are used to detect seamounts as local sinks. Other techniques that have been applied to similar problems in other fields include image classification using fuzzy logic [e.g., *Drăguţ and Blaschke*, 2006] and various rule-based approaches [e.g., *Iwahashi and Pike*, 2007].

[4] The limited coverage of ship surveys and the low resolution of satellite measurements remain a major challenge for any seamount census. Unless conditions such as water depth and average sea state are favorable, seamounts less than $\sim$ 1 km tall are unlikely to be detected; they are thus believed to be severely under-represented in existing catalogs [*Hillier and Watts*, 2007; *Kim and Wessel*, 2011]. Even for larger seamounts, catalogs may be incomplete: *Hillier and Watts* [2007] suggest that up to 60% of seamounts over 1 km in height may yet be undiscovered.

[5] Most search algorithms rely on a simple *a priori* model of an "ideal" seamount, often of a Gaussian or polynomial form in cross section. This is unlikely to be a good model in all cases and risks introducing systematic biases into results. In addition, many areas—including much of the Atlantic Ocean—are characterized by closely spaced fracture zones flanked by topographic highs, and by strong "tectonic fabric": rough, low amplitude topography formed by fundamentally different processes to seamounts [e.g.,*Macdonald et al.*, 1996]. Simple models often struggle to distinguish between these different morphologies, resulting in many false "seamounts" being identified.

[6] One potential solution lies in the use of learning algorithms, such as neural networks, designed to identify patterns and similarities in a set of "training data": a relatively small number of representative examples selected manually by the researcher. Once trained, the network can be used to systematically assess how similar new examples are to those in the training set. Depending on the intended application, one can then identify new "matching" features or delineate regions worthy of further inspection.

[7] A wide range of approaches are encompassed by the terms "learning algorithm" or "neural network," which remain areas of active research. A comprehensive introduction can be found in *Bishop* [1995] or *Mackay* [2003]; reviews of geophysical applications include *van der Baan and Jutten* [2000] and *Sandham and Leggett* [2004]. Within geomorphology, neural networks have been used for drainage modeling [e.g., *Zhang and Govindaraju*, 2003], landslide risk assessment [e.g., *Ermini et al*., 2005], and land use classification [e.g., *Ehsani and Quiel*, 2009]. However,

Additional supporting information may be found in the online version of this article.

[1]Department of Earth Sciences, Universiteit Utrecht, Utrecht, Netherlands.
[2]Department of Earth Sciences, University of Oxford, Oxford, UK.

Corresponding author: A. P. Valentine, Department of Earth Sciences, Universiteit Utrecht, PO Box 80021, NL-3508 TA Utrecht, Netherlands. (andrew@geo.uu.nl)


0094-8276/13/10.1002/grl.50615

we are not aware of any previous applications to systematic search for and cataloging of landforms.

[8] Many different types of neural network might be applied to the task of detecting seamounts, and there is no single "correct" approach. Here we use an "autoencoder", designed to find low-dimension representations of complex data sets [*Hinton and Salakhutdinov*, 2006]. By reducing the dimensionality, we attempt to find a (mathematical) space in which data of interest can be represented, but other data cannot [*Valentine and Trampert*, 2012]. If previously unseen examples can then be represented accurately, they are likely to be similar to those in the training set: quantifying the accuracy of representation allows assessment of the degree of similarity. Here we describe our method and demonstrate it for $10° \times 10°$ regions from the Pacific and Atlantic Oceans.

## 2. Detection Algorithm

[9] Our approach is based on the analysis of individual "patches"—spatial windows—of topography; we test each to assess the likelihood that it is centered upon a feature of interest, here a seamount. In designing this system, we must make many decisions: How large should each patch be? What preprocessing should be performed? Is postprocessing required? In what follows, we present one set of choices, which have been found to give satisfactory performance. However, much further investigation is required to fully understand their impact on results, and it is almost certain that the current setup is not optimal. In any case, the definition of "optimal" will vary between applications, with different choices being required. However, our experience suggests that performance is not particularly sensitive to moderate changes in the various parameters and that satisfactory results can be achieved with a wide range of choices.

### 2.1. Training Data Set

[10] We first need to construct a training set of exemplar seamounts, from which the network will learn to assimilate and recognize the characteristics that define these features. For the purposes of this case study, we make use of the fact that catalogs already exist, and extract random examples from the global list compiled by *Kim and Wessel* [2011]. We visually inspect the patches centered on each of these locations and select only those that have a clear seamount in their central region. In this manner, we obtain 1000 examples for our training set and a further 1000 independent examples to assess and monitor network performance during training. This ensures that information learned by the network is general (applicable to unseen examples) and not specific to the training set.

[11] By building our training set in this way, we clearly risk biasing our analysis toward the data selection and processing choices made by *Kim and Wessel* [2011], and the training set will exclude any classes of seamount systematically excluded from their analysis. This must be borne in mind when interpreting results; ideally, the training data set would be constructed in a manner that is independent of any *a priori* assumptions.

[12] We also note that the training set used here is large: this ensures good performance but may be impractical for certain studies. Further work is necessary to explore performance for smaller training sets and develop strategies for reducing the amount of hand-selection required—for example, it might be possible to devise a "bootstrapping" approach, where an initial small hand-selected data set is then used to construct a larger training set.

### 2.2. Data Processing

[13] For each seamount in our training set, we wish to convert the bathymetry patch to a vector, $\mathbf{s}$, of processed surface heights. In order that sensitivity to size and shape remains uniform across the globe, these surface heights should correspond to points that are equally spaced in kilometers within the patch. To keep computational costs reasonable, we wish to limit the length of the vector while ensuring that individual patches span sufficient area to enable all relevant features to be detected. We therefore down-sample the raw data, with a trade-off between patch size, sampling density, and computational complexity. Each patch is extracted from the global seafloor topography data set of *Smith and Sandwell* [1997, v.14.1 released 2011], which has 1 arc min resolution and combines satellite bathymetry (inferred from variations in the Earth's gravitational field) with high-resolution ship-based bathymetry, where available. We convert a 150 km × 150 km grid to 64 × 64 points, i.e., $\mathbf{s}$ contains 4096 elements. This resampling implies there is little benefit for us in extracting data from higher-resolution grids such as SRTM30_PLUS [*Becker et al.*, 2009].

[14] To obtain a single vector, $\mathbf{s}$, containing data for the patch centered on latitude/longitude coordinates $(\theta_c, \phi_c)$, we do the following:

[15] 1. Project the raw data (an oversized region, to allow for edge effects) using a Lambert Azimuthal projection, which is equal-area, with $(\theta_c, \phi_c)$ as the projection center;

[16] 2. Filter the data using a 10th-order Butterworth low-pass filter, to prevent aliasing during downsampling, using a frequency cutoff 2.5× Nyquist's frequency;

[17] 3. Resample the data so that the 150 km × 150 km region is described by 64 × 64 uniformly spaced points;

[18] 4. Scale each point in the patch by a factor $\alpha(r) = (1 - ar)^2$, with $a = 6/750$ and where $r$ represents the distance from $(\theta_c, \phi_c)$ in kilometers, so that structure toward the edges are downweighted relative to the center;

[19] 5. "Unwrap" the two-dimensional patch into a single vector by traversing lines of latitude;

[20] 6. Subtract the median value and scale the vector by its maximum absolute value, so that values lie in the range $[-1, 1]$.

Some examples of raw bathymetry and corresponding processed data as input to the network may be seen in Figure 1.

### 2.3. The Autoencoder

[21] Each patch is processed to yield a 4096-element topography vector, $\mathbf{s_i}$. However, since all patches that are centered on seamounts inherently have a degree of similarity, we postulate that they can be represented using significantly less than 4096 independent parameters. If this is the case, a mathematical operation exists that "encodes" each $\mathbf{s_i}$, allowing it to be represented by an $M$-element vector $\mathbf{t_i}$

$$\mathbf{t_i} = \underset{4096 \rightarrow M}{\text{enc}}\, \mathbf{s_i}\,, \tag{1}$$

with a corresponding "decoding" operation

$$\mathbf{s_i'} = \underset{M \rightarrow 4096}{\text{dec}}\, \mathbf{t_i}\,. \tag{2}$$

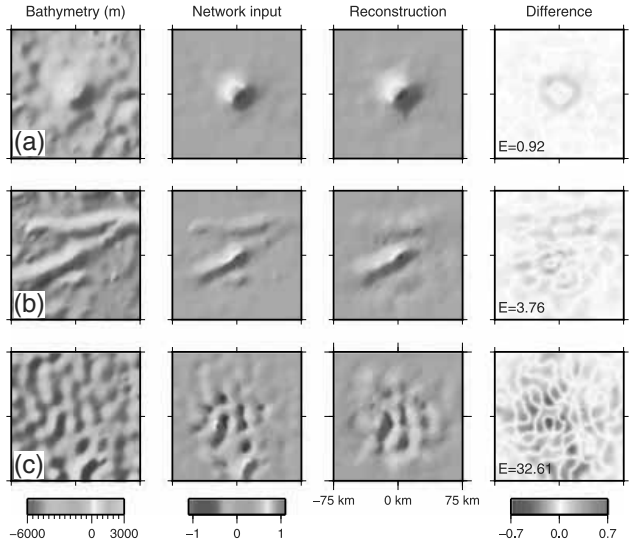


**Figure 1.** Examples of data processing. The 150 km × 150 km regions of bathymetric data [*Smith and Sandwell*, 1997] are weighted and rescaled, before being presented to a neural network trained to only represent seamount-like structure. The difference between network inputs and reconstruction can be quantified by the reconstruction error, $E$ (equation (4)), and used to assess whether a patch is centered on a seamount. Figure 1a is an archetypal isolated seamount, whereas Figure 1b is an elongated ridge but may be regarded as seamount-like in origin. Figure 1c represents a region free from volcanic features. These examples are drawn from the Pacific demonstration region (see section 3) and are not used by the network during training.

These operations can be implemented by an autoencoder network—essentially, an encoder-decoder pair—as discussed in *Valentine and Trampert* [2012]; the notation used in that paper is consistent with the notation here. In the autoencoder, the data $\mathbf{s_i}$ is passed through layers with successively fewer parameters to obtain an encoding, then passed through layers with successively more to yield the reconstruction $\mathbf{s_i'}$. In the present case, we reduce or increase the number of parameters by a factor of 2 between successive layers and choose $M = 64$.

[22] If our encoding captures the important features of the data, $\mathbf{s_i'}$ should closely resemble $\mathbf{s_i}$. The network behavior is governed by a large number of adjustable weights, and we train the network by optimizing these to minimize the average "reconstruction error"

$$\bar{E} = \frac{1}{2Q} \sum_{i=1}^{Q} \left| \mathbf{s_i'} - \mathbf{s_i} \right|^2 , \tag{3}$$

obtained across a training set of $Q$ example patches all centered on seamounts (here $Q = 1000$). Full details of the training algorithm are given in *Valentine and Trampert* [2012]. The result is a seamount-optimized network that is unable to accurately represent non-seamount-centered patches: the 64 elements of the encoding cannot transmit the information needed to reconstruct other topographies. We can therefore use the reconstruction error of a patch,

$$E = \frac{1}{2} \left| \mathbf{s'} - \mathbf{s} \right|^2 , \tag{4}$$

to assess the likelihood that it is centered upon a seamount. This is a common formulation in neural network training algorithms, and full details are given in *Valentine and Trampert* [2012]. While squared-difference measures may be sensitive to outliers, this is unlikely to pose difficulties in the present case: by virtue of the processing carried out in constructing the global bathymetry grid, as well as our preprocessing, we expect patches and their reconstructions to be relatively smooth. Some examples of the reconstructions produced by a network trained in this way can be seen in Figure 1, along with their reconstruction error.

### 2.4. A Systematic Search

[23] Once the network has been successfully trained, we wish to systematically search a region for seamount-like structures. At each point on a 1 min grid, we extract a patch as described in section 2.2, with one exception: when rescaling the patch in step 6, we generate a regional grid of maximum absolute patch values, and then low-pass filter this (10th-order Butterworth, wavelength 20 km). The filtered values are used to rescale patches: this helps preserve relationships between adjacent sites, and reduces artifacts. Each patch is encoded and reconstructed by the network, and the patch reconstruction error is computed according to equation (4). We can therefore assemble a spatial grid of reconstruction error; the lower the error at a given point, the higher the degree of similarity between the local topography and the training examples—in this case, seamounts.

[24] This grid of reconstruction error, $E$, is the most fundamental result we obtain, and for certain applications, it may be useful in its own right. However, it is often necessary to identify discrete locations corresponding to probable seamounts, which corresponds to picking local minima in the error surface. However, the error surface contains small-scale fluctuations which must be regarded as noise, and we need to determine whether a particular minimum is "significant." Many different approaches and tests could be used here, and for present purposes, we adopt a relatively simple method.

[25] The $E$ grid is calculated over a region enlarged by 0.5° to mitigate edge effects. We then reduce the short wavelength noise by applying a Butterworth low-pass filter with wavelength $\lambda$. We require a location to have a reconstruction error $E < E_{\max}$ to be considered further; within areas meeting this criterion, it is straightforward to identify all local minima by comparing sites to their eight nearest neighbors. We also compute the relative depth of each minimum, by "growing" a contour from the minimum until it encompasses at least one other local minimum. If this depth is less than $D_{\min}$, we discard the original minimum unless it is lower than the adjacent minimum. Typically, removed points correspond to minor depressions on the flanks of deeper minima.

[26] In processing each patch, the data is rescaled to ensure all network inputs lie in the range [–1, 1] (step 6, section 2.2). This is necessary to ensure that the network has uniform sensitivity; however, it prevents the network, and hence the reconstruction error, from taking the absolute height of features into account directly. We therefore perform an additional test and determine the local "prominence"

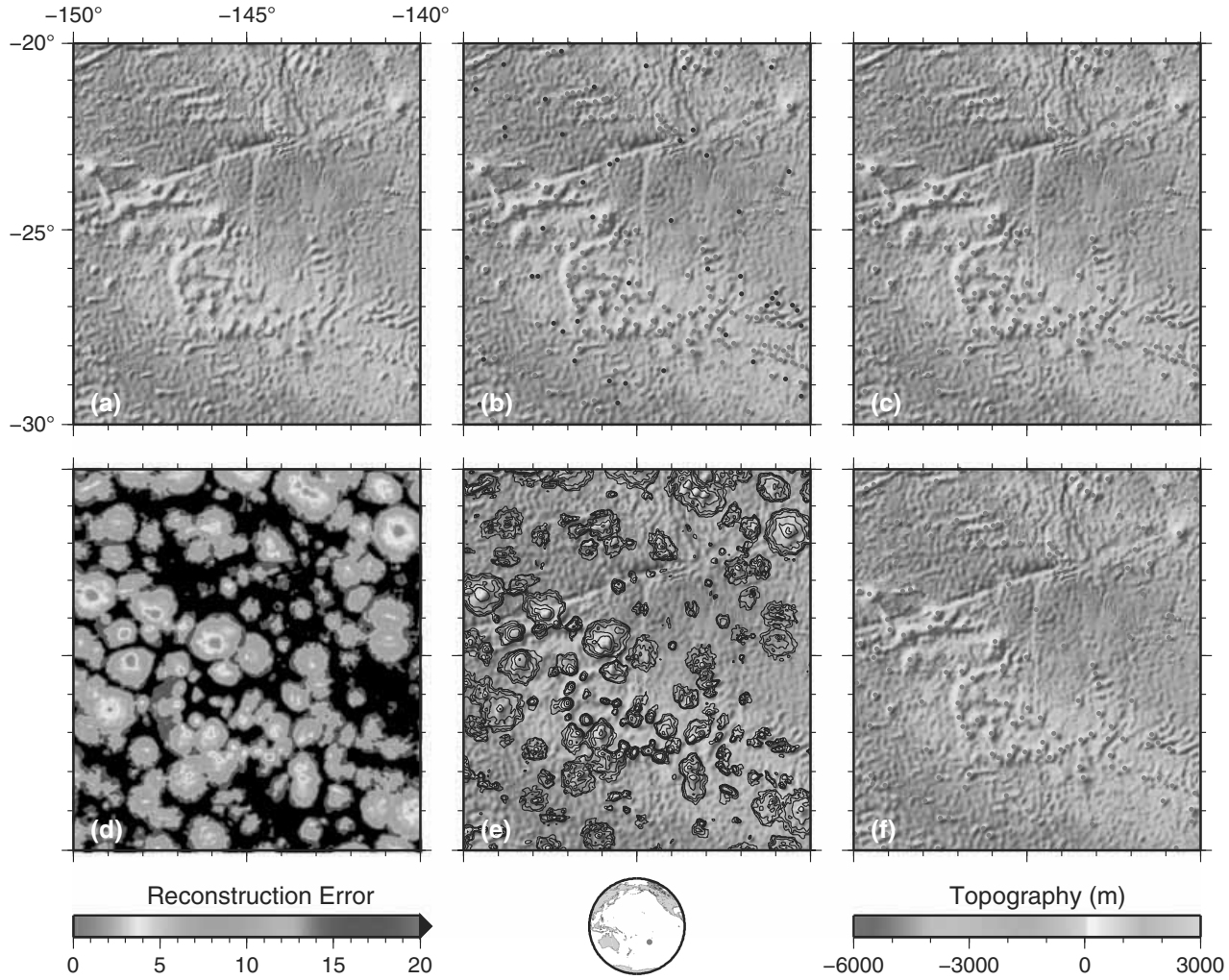


**Figure 2.** Demonstration, 10° × 10° region, Pacific Ocean (see map, inset). (a) Bathymetric data [*Smith and Sandwell*, 1997]; numerous seamounts are clearly visible. (b) Our handpicked seamounts—definite seamounts in bright red, possible seamounts in dark red. (c) Seamounts as cataloged by *Kim and Wessel* [2011]. (d) Reconstruction error computed by a network trained to only represent seamounts. (e) Contours $(1, 3, \ldots, 11)$ of reconstruction error overlain on bathymetry. (f) Discrete seamounts identified by picking local minima of reconstruction error.

by high-pass filtering bathymetry with a 10th-order Butterworth filter, with wavelength $\lambda_P$. We require all candidate seamount locations to have a local prominence (height after filtering) of at least $P$.

## 3. Demonstration

[27] To demonstrate the potential of this approach, we show performance on two 10° × 10° regions. The first, in the Pacific, has been used in other investigations, such as *Wessel and Lyons* [1997]. The second, from the Atlantic, lies in the vicinity of the Azores; this general area has featured in studies such as those by *Morato et al.* [2008] and *Yesson et al.* [2011].

[28] From Figures 2d, 2e, 3d, and 3e, it is apparent that the $E$ grid shows an excellent correlation with seamount-like topography, and this grid may have direct applications: for example, calculating the volume of seamounts or the fraction of the ocean they cover. However, as discussed above, many applications require us to obtain a list of discrete seamounts by isolating minima in the error surface. In order to assess the success of this process, we compare automatic selection with handpicked seamounts in the two regions.

[29] Our manual identification of seamounts in the demonstration regions are shown in Figures 2b and 3b. We divide these into "definite" seamounts—features that are clearly seamounts and should be identified by a successful algorithm—and "possible" seamounts, where the data are ambiguous. An algorithm should neither be rewarded for picking these "possible" features, nor penalized for missing them. Obviously, such hand-selection is somewhat subjective, as no two experts are likely to make identical choices. Many features, though, are relatively unambiguous, and we do not believe that minor differences in selection criteria would significantly affect the overall analysis of the results. Note that similar subjective choices were made in hand-cleaning the training data set: fundamental and widespread differences there would be reflected in the error surface.

[30] We assess the accuracy of a picking method using two metrics: success rate ($SR$), the percentage of "definite"

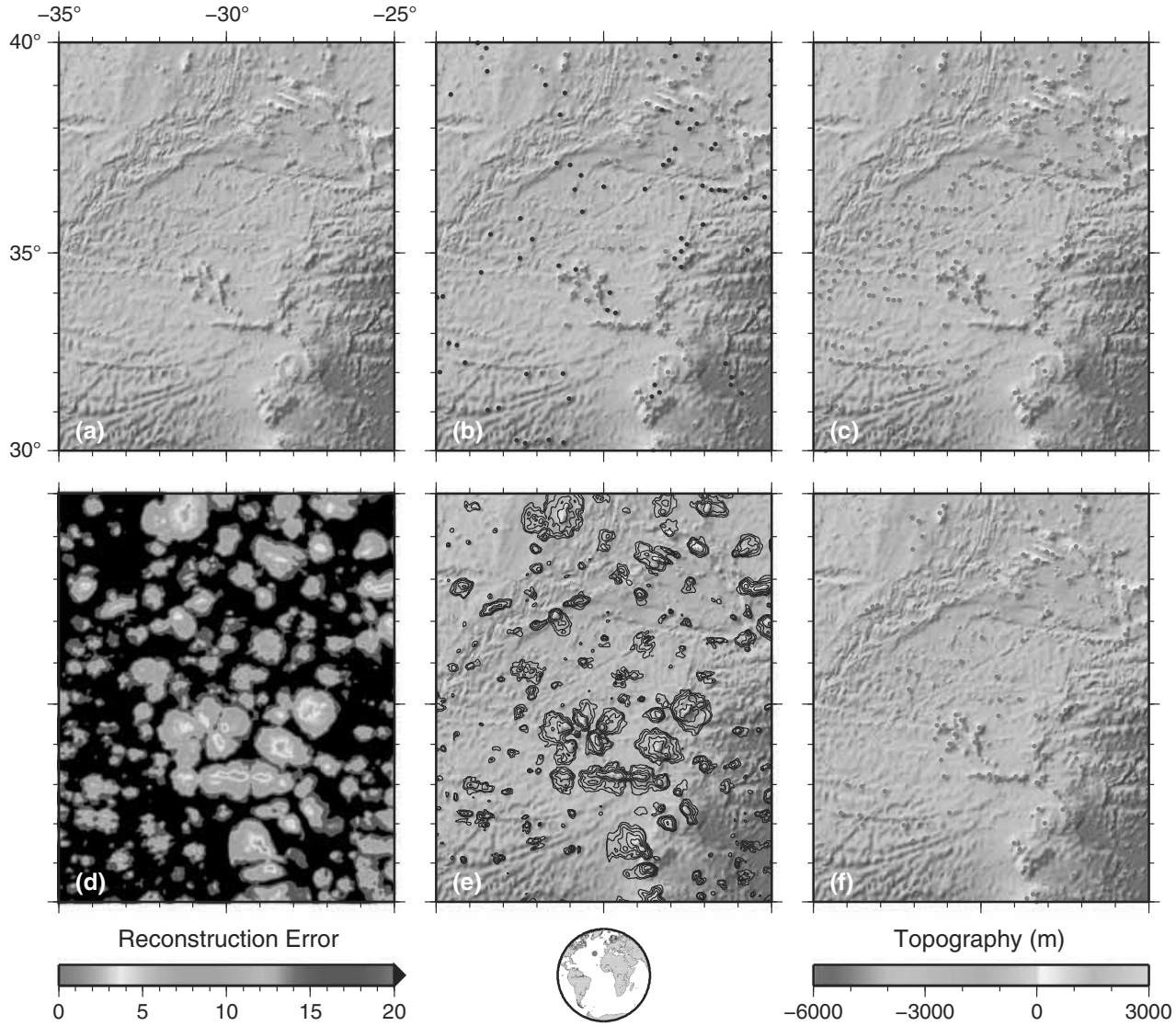


**Figure 3.** Demonstration, 10° × 10° region, Atlantic Ocean. As in Figure 2.

seamounts found, and false positive rate (*FPR*), the percentage of automated picks that do not correspond to either "definite" or "possible" seamounts; the numbers of "matches" used in these calculations are provided in Table S1. Obviously, results depend on the choices made regarding the various parameters $\lambda$, $E_{\max}$, $D_{\min}$, etc. In Figures 2f and 3f, we show one example, using a third-order Butterworth filter of wavelength 10 km, selecting minima with $E < 7$ and depth greater than 0.1, and requiring a prominence of 300 m (with $\lambda_P$ = 100 km). We obtain *SR*/*FPR* = 73.2%/5.2% for the Pacific region and 67.1%/33.7% for the Atlantic; the higher *FPR* in the Atlantic is largely due to the challenges posed by the much rougher tectonic fabric. For comparison, computing the equivalent statistics between our hand-picks and the results of *Kim and Wessel* [2011] yields 72.8%/4.9% and 70.9%/61.9% for the Pacific and Atlantic regions, respectively, when all their locations are used. If we exclude locations for which they report a height under 300 m (corresponding to our prominence test), the Pacific results are unchanged; their Atlantic rates become 68.4%/52.4%. In determining whether a given hand-pick matches a particular automatic pick, we accept a separation of up to 25 km in order to accommodate large seamounts and volcanic ridges, but the vast majority of matches are far closer: the median distance across both "possible" and "definite" seamounts is 3.49 km and 3.83 km for the Pacific and Atlantic regions, respectively, for the autoencoder picks.

[31] Our results therefore compare favorably with *Kim and Wessel* [2011]; the success rates are similar, but the false positive rate in the Atlantic is much lower, to a large extent because the autoencoder has a substantially better performance over the rough tectonic fabric in the southern part of the region. However, the majority of the remaining false positives still come from abyssal hills, ridges flanking fracture zones, and the topography of the spreading ridge itself. Most false negatives—seamounts that the algorithm fails to find—are smaller features; using a smaller window with the full resolution of the bathymetry might help detect these. Alternatively, it might be beneficial to use different data types, such as vertical gravity gradient (VGG), either in

place of, or in combination with, the bathymetry. The VGG has more power at short wavelengths than satellite-derived bathymetry, although it is substantially less sensitive than ship-based bathymetry, which, where available, is incorporated into the global bathymetric grid. Thus, use of VGG may aid in the detection of some smaller features currently missed by the algorithm but may also increase the number of false positives due to features such as fracture zone ridges, which are also short wavelength features [as seen in *Kim and Wessel*, 2011].

[32] As with most automatic classification or picking algorithms, there is a strong trade-off between success rate and false positive rate. With different choices of parameters for the picking process, it is possible to raise the success rate over 80%, at the cost of an increased false positive rate, e.g., *SR*/*FPR* = 87.2%/27.4% for the Pacific and 81.0%/58.7% for the Atlantic. At the other extreme, for *FPR* < 1%, it is possible to achieve *SR* = 63.1% and 37.2% for the Pacific and the Atlantic, respectively, albeit using different parameters for each (see Table S1). Indeed, best results may well require parameters to vary between regions. It might be possible, for example, to construct a scheme that varies the choice of processing parameters based on spreading rate or roughness, thus automatically tuning the algorithm to the expected nature of the local topography.

[33] This paper seeks to demonstrate that neural networks may be applied to the cataloging of geomorphological features, of which seamounts are only one example. The "correct" balance between *FPR* and *SR* will depend on the application, and we do not attempt to give a single best choice nor analyze the detailed implications of differences between our results and other studies. We also note that we have analyzed performance only for the two regions presented here, and results for other regions may differ. One of the strengths of our approach is that it is completely automated: unlike previous studies, we have performed no hand-cleaning nor introduced special exclusion zones, e.g., near the ridge axis or along fracture zones, although such strategies could be employed if required. Given the raw *E* grid, a user can easily adjust the picking parameters to the levels best suited for their application and region of interest, or design new heuristics for extracting discrete picks from the *E* data. This offers additional adaptability and versatility compared with traditional methods.

## 4. Conclusions

[34] We have successfully developed a learning algorithm approach for identifying seamounts in global bathymetry data, with results that compare favorably with those from traditional methods. This is particularly evident in our Atlantic region, where rough tectonic fabric makes seamount identification particularly challenging; the autoencoder method yields substantially fewer false positives as its complex pattern recognition capabilities enable it to better distinguish between "background" tectonic signatures and seamounts arising from excess volcanism. The complete automation of the method of extracting discrete probable seamounts from the continuous *E* grid adds power and flexibility; different sets of picks can easily be extracted for different purposes. Identification of smaller seamounts remains limited by the limited coverage of high-resolution data; however, where such data are available, a similar framework could be deployed, optimized for smaller features if necessary. Further work is needed to fully develop the method and exploit its full power: refinement of the picking method, analysis of the information contained in the encoding data itself, and development of methods for smaller training sets. In principle, however, little of the method here is specific to our seamount case study; the technique is highly versatile and could be applied to identifying and analyzing many different types of geomorphic features.

[35] **Acknowledgments.** We are grateful to the developers of the Generic Mapping Tools [*Wessel and Smith*, 1991], which have simplified our work immeasurably. We thank Tony Watts for many informative discussions on marine geomorphology and the distinctions between abyssal hills, seamounts, and volcanic ridges; Paul Käufl for conversations which influenced our computational implementation of the picking algorithms; and Theo van Zessen for looking after The Stig. Reviews from Seung-Sep Kim and an anonymous reviewer helped improve this manuscript. APV is supported by the Netherlands Organisation for Scientific Research (NWO), through *Topsubsidie* 854.10.002. LMK is supported by the UK Natural Environment Research Council through grant NE/I026839/1 and UKIODP grant NE/J011401/1.

[36] The Editor thanks Seung-Sep Kim and Joseph Becker for their assistance in evaluating this paper.

## References

Becker, J., et al. (2009), Global bathymetry and elevation data at 30 arc seconds resolution: SRTM30_PLUS, *Mar. Geod.*, *32*, 355–371.

Bishop, C. (1995), *Neural Networks for Pattern Recognition*, Oxford Univ. Press, Oxford, U. K.

Craig, C., and D. Sandwell (1988), Global distribution of seamounts from Seasat profiles, *J. Geophys. Res.*, *93*, 10,408–10,420.

Drăguţ, L., and T. Blaschke (2006), Automated classification of landform elements using object-based image analysis, *Geomorphology*, *81*, 330–344.

Ehsani, A., and F. Quiel (2009), A semi-automatic method for analysis of landscape elements using Shuttle Radar Topography Mission and Landsat ETM+ data, *Comput. Geosci.*, *35*, 373–389.

Ermini, L., F. Catani, and N. Casagli (2005), Artificial neural networks applied to landslide susceptibility assessment, *Geomorphology*, *66*, 327–343.

Hillier, J. (2008), Seamount detection and isolation with a modified wavelet transform, *Basin Res.*, *20*, 555–573.

Hillier, J., and A. Watts (2007), Global distribution of seamounts from ship-track bathymetry data, *Geophys. Res. Lett.*, *34*, L13304, doi:10.1029/2007GL029874.

Hinton, G., and R. Salakhutdinov (2006), Reducing the dimensionality of data with neural networks, *Science*, *313*, 504–507.

Iwahashi, J., and R. Pike (2007), Automated classifications of topography from DEMs by an unsupervised nested-means algorithm and a three-part geometric signature, *Geomorphology*, *86*, 409–440.

Kim, S.-S., and P. Wessel (2011), New global seamount census from altimetry-derived gravity data, *Geophys. J. Int.*, *186*, 615–631.

Kitchingman, A., and S. Lai (2004), Inferences on potential seamount locations from mid-resolution bathymetric data, in *Seamounts: Biodiversity and Fisheries, Fisheries Centre Research Report*, edited by T. Morato, and D. Pauly, pp. 7–12, vol. 12, Univ. of B. C., Vancouver.

Macdonald, K., P. Fox, R. Alexander, R. Pockalny, and P. Gente (1996), Volcanic growth faults and the origin of Pacific abyssal hills, *Nature*, *380*, 125–129.

Mackay, D. (2003), *Information Theory, Inference and Learning Algorithms*, Cambridge Univ. Press, Cambridge, U. K.

Menard, H. (1959), Geology of the Pacific seafloor, *Experimentia*, *15*, 205–244.

Morato, T., M. Machete, A. Kitchingham, F. Tempura, S. Lai, G. Menezes, T. Pitcher, and R. Santos (2008), Abundance and distribution of seamounts in the Azores, *Mar. Ecol. Prog. Ser.*, *357*, 17–21.

Sandham, W., and M. Leggett (2004), *Geophysical Applications of Artificial Neural Networks and Fuzzy Logic*, Kluwer Acad., Dordecht, Netherlands.

Smith, D., and T. Jordan (1988), Seamount statistics in the Pacific Ocean, *J. Geophys. Res.*, *93*, 2899–2918.

Smith, W., and D. Sandwell (1997), Global seafloor topography from satellite altimetry and ship depth soundings, *Science*, *277*, 1957–1962.

Valentine, A., and J. Trampert (2012), Data space reduction, quality assessment and searching of seismograms: Autoencoder networks for waveform data, *Geophys. J. Int.*, *189*, 1183–1202.

van der Baan, M., and C. Jutten (2000), Neural networks in geophysical applications, *Geophysics*, *65*, 1032–1047.

Wessel, P., and S. Lyons (1997), Distribution of large Pacific seamounts from Geosat/ERS-1: Implications for the history of intraplate volcanism, *J. Geophys. Res.*, *102*, 22,450–22,475.

Wessel, P., and W. Smith (1991), Free software helps map and display data, *Eos Trans. AGU*, *72*, 441–446.

Yesson, C., M. Clark, M. Taylor, and A. Rogers (2011), The global distribution of seamounts based on 30 arc seconds bathymetry data, *Deep Sea Res. I*, *58*, 442–453.

Zhang, B., and R. Govindaraju (2003), Geomorphology-based artificial neural networks (GANNs) for estimation of direct runoff over watersheds, *J. Hydrol.*, *273*, 18–34.